

Ostermeier, Christian; Carstensen, Claus H.; Prenzel, Manfred; Geiser, Helmut  
**Kooperative unterrichtsbezogene Qualitätsentwicklung in Netzwerken.  
Ausgangsbedingungen für die Implementation im  
BLK-Modellversuchsprogramm SINUS**

*Unterrichtswissenschaft 32 (2004) 3, S. 215-237*



Quellenangabe/ Reference:

Ostermeier, Christian; Carstensen, Claus H.; Prenzel, Manfred; Geiser, Helmut: Kooperative unterrichtsbezogene Qualitätsentwicklung in Netzwerken. Ausgangsbedingungen für die Implementation im BLK-Modellversuchsprogramm SINUS - In: Unterrichtswissenschaft 32 (2004) 3, S. 215-237 - URN: urn:nbn:de:01111-opus-58147 - DOI: 10.25656/01:5814

<https://nbn-resolving.org/urn:nbn:de:01111-opus-58147>

<https://doi.org/10.25656/01:5814>

in Kooperation mit / in cooperation with:

**BELTZ JUVENTA**

<http://www.juventa.de>

**Nutzungsbedingungen**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, veröffentlichen oder widernatürlich nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

**Terms of use**

We grant a non-exclusive, non-transferable, individual and limited right to using this document.  
This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

**Kontakt / Contact:**

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Digitalisiert

---

# Unterrichtswissenschaft

Zeitschrift für Lernforschung

32. Jahrgang / 2004 / Heft 3

---

## Thema

### Implementationsforschung

Verantwortliche(r) Herausgeber(in)

Cornelia Gräsel, Peter Strittmatter

*Cornelia Gräsel, Peter Strittmatter*

Einführung..... 194

*Cornelia Gräsel, Ilka Parchmann*

Implementationsforschung – oder: der steinige Weg,

Unterricht zu verändern..... 196

*Christian Ostermeier, Claus H. Carstensen,*

*Manfred Prenzel, Helmut Geiser*

Kooperative unterrichtsbezogene Qualitätsentwicklung in

Netzwerken: Ausgangsbedingungen für die Implementation

im BLK-Modellversuchsprogramm SINUS..... 215

*Anja Fey, Cornelia Gräsel, Thomas Puhl, Ilka Parchmann*

Implementation einer kontextorientierten Unterrichtskonzeption

für den Chemieunterricht..... 238

*Robin Stark*

Eine integrative Forschungsstrategie zur anwendungsbezogenen

Generierung relevanten wissenschaftlichen Wissens in der

Lehr-Lern-Forschung ..... 257

## Allgemeiner Teil

*Daniel Preckel*

Problembasiertes Lernen: Löst es die Probleme der traditionellen

Instruktion?..... 274

## Kooperative unterrichtsbezogene Qualitätsentwicklung in Netzwerken

Ausgangsbedingungen für die Implementation im  
BLK-Modellversuchsprogramm SINUS

Developing Instructional Quality in School-Networks. Baseline  
Conditions for the Implementation in a Pilot Program to Increase  
the Efficiency of Mathematics and Science Instruction

---

*Implementationsforschung wird im folgenden Beitrag unter dem Aspekt bedeutsamer Ausgangsbedingungen an den von einer Innovation betroffenen Institutionen betrachtet. Im Blickpunkt steht der Qualitätsentwicklungsansatz des BLK-Programms zur „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“ (SINUS). Aus Sicht der Implementationsforschung gilt es zu klären, inwieweit theoretisch bedeutsame Bedingungen die Umsetzung des Programms beeinflussen und welche dieser Bedingungen sich förderlich oder hinderlich auswirken. Um Implementationsforschung mit dieser Fragestellung betreiben zu können, müssen systematische Informationen über relevante Bedingungsbereiche vorliegen und zwar vor Einsetzen einer Programmwirkung. Im Beitrag werden bedeutsame Ausgangsbedingungen an den beteiligten Programmschulen untersucht. Dabei wird geprüft, ob die beteiligten Schulen eine besondere Selektion in Bezug auf unterrichts- und schulbezogene Ausgangslagen darstellen. Es werden vier Bereiche betrachtet: Familiärer Hintergrund der Schülerinnen und Schüler, motivationale Orientierungen und Präferenzen im mathematisch-naturwissenschaftlichen Bereich, Schulklima und Kompetenzen in Mathematik und Naturwissenschaften. Bei der Erhebung der Ausgangslage wurden 136 SINUS-Schulen mit national entwickelten PISA-Instrumenten getestet. Die Ergebnisse können daher mit Daten einer für Deutschland repräsentativen Stichprobe (PISA/E 2000) verglichen werden. Die Ergebnisse zeigen, dass sich das Programm nicht systematisch auf eine Selektion besonderer Schulen bezieht. Sie werden zum einen in Hinblick auf die Implementationsforschung, zum anderen in ihrer Bedeutung für die allgemeine Übertragbarkeit der Implementation an weiteren Schulen diskutiert.*

*In the article implementation research is treated in the light of relevant conditions at schools affecting an innovation. The quality development approach of the pilot program to increase the efficiency of mathematics and science instruction (SINUS) serves as an example to investigate this question. From an implementation research perspective it is important to examine which conditions foster or hamper the implementation of the approach. To study these questions systematic data concerning relevant conditions at the beginning of the program have to be available. Baseline data from the participating schools will be reported. Whether or not the participating schools have been recruited from a special selection of (e. g. excellent) schools will also be reported. Four aspects will be looked at: family background of the students, motivational variables with regard to mathematics and science teaching, school climate and competencies in mathematics and science. 136 program schools took part in the baseline study using nationally developed German PISA instruments. The results can, therefore, be compared with a representative group of schools (PISA/E 2000, the national representative extension of the German PISA sample). The study provides information on relevant baseline conditions in the participating schools that is needed as a prerequisite for carrying out implementation research. Results show that the program does not draw on a special selection of schools. These results are discussed in the light of implementation research, and in their meaning for generalizing the implementation for a larger group of schools.*

Mit der Einrichtung von Interventionsprogrammen im Bereich der Schule (Prenzel & Achtenhagen, 2000) ist die Frage nach der Wirksamkeit dieser Projekte verbunden. Deutlich wird dies beispielsweise in den Richtlinien zur Neuordnung der Modellversuchsförderung der Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (BLK), in denen die systematische Evaluation der Fördermaßnahmen verlangt wird (Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung, 1997b).

Eng verknüpft mit der Frage nach der Wirkung der Programme ist das Problem der Implementation: Wie wurde die Innovation in den schulischen Alltag umgesetzt und welche Rahmenbedingungen befördern oder behindern eine erfolgreiche Realisierung der Konzeption? Einflussfaktoren für eine gelingende Implementation können dabei verschiedenen Ebenen zugeschrieben werden. In Implementationsmodellen (Blumenfeld, Fishman, Krajcik & Marx, 2000; Euler & Sloane, 1998) werden beispielsweise drei Ebenen unterschieden: Merkmale der beteiligten Lehrkräfte, Merkmale der Schulen sowie curriculare und schulorganisatorische Rahmenbedingungen.

Der vorliegende Beitrag behandelt Voraussetzungen an Schulen, die sich auf die Umsetzung einer innovativen Konzeption auswirken. Im Blickpunkt steht dabei das BLK-Modellversuchsprogramm zur „Steigerung der Effi-

zienz des mathematisch-naturwissenschaftlichen Unterrichts“ (SINUS). Das Programm zielt auf eine Verbesserung des mathematisch-naturwissenschaftlichen Unterrichts in der Sekundarstufe ab. Die Weiterentwicklung des Unterrichts erfolgt auf der Grundlage verfügbarer fachdidaktischer und allgemeindidaktischer Konzeptionen und Beispiele. Dem Programm liegt ein Implementationsansatz zu Grunde, der sich durch folgende Merkmale auszeichnet (Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung, 1997a; Prenzel, 2000; Prenzel & Ostermeier, 2003):

- Die Unterrichtsentwicklung erfolgt schulnah und orientiert sich an bekannten Problembereichen des mathematisch-naturwissenschaftlichen Unterrichts. Ausgangspunkt und Rahmen für die Weiterentwicklung ist ein System von elf Modulen, die sich auf die Problembereiche beziehen und so zugeschnitten sind, dass ein Einstieg in, sowie ein Austausch über die Arbeit erleichtert wird.
- Am Programm beteiligte Lehrkräfte werden über die Arbeit an ihren Schulen mit Verfahren der kooperativen Qualitätssicherung und Qualitätsentwicklung vertraut gemacht. Unterstützt durch wissenschaftliche Anregungen (z.B. Fortbildungen, Handreichungen, Beratungen) sollen diese Prozesse in den Schulen langfristig etabliert werden.
- Kooperation gilt als zentrales Leitprinzip im Programm. Gefordert und unterstützt wird die Zusammenarbeit zwischen Lehrkräften an der Schule, zwischen Lehrkräften verschiedener am Programm beteiligter Schulen in regionalen und überregionalen Netzwerken, sowie zwischen Schulpraxis und allgemeiner und fachspezifischer Lehr-Lernforschung.
- Die Zusammenarbeit wird an der Schule selbst, sowie auf der Ebene der Schulnetzwerke durch Koordinatorinnen und Koordinatoren unterstützt. Die Unterstützung beinhaltet Anregungen und Hilfen für die Qualitätsentwicklung, Organisation des Austauschs und Dokumentation der Ergebnisse und Prozesse. Auch werden kollegiale Fortbildungen auf Schul- und Netzwerkebene angeboten.
- Ein zentraler Programmträger hält Handreichungen sowie Materialien zu den Modulen bereit und stellt Fortbildungs- und Beratungsangebote zur Verfügung. Außerdem besteht ein zentraler Server zum Informationsmanagement und zur Unterstützung des Austauschs.
- Begleitforschung als bedeutsamer Teil der Programmkonzeption umfasst Erhebungen zu formativer Evaluation der Unterstützungsleistungen und zur Akzeptanz des Programms, sowie Erhebungen im Rahmen der summativen Evaluation. Die unterschiedlichen Erhebungsstränge erlauben außerdem die Beantwortung von spezifischen Fragen im Rahmen der Implementationsforschung.

Die beteiligten Lehrerinnen und Lehrern nehmen im Programm eine sehr aktive Rolle ein. Als Expertinnen und Experten für Lehren und Lernen reflektieren sie ihren eigenen Unterricht und entwickeln diesen weiter (Schön, 1987). Einen Ansatzpunkt und Rahmen für die Unterrichtsentwicklung bil-

den dabei die oben genannten Module. Gleichzeitig steht den Lehrkräften, wie beschrieben, ein vielfältiges Unterstützungssystem zur Verfügung.

Die genannten Leitlinien beschreiben insgesamt einen Implementationsansatz, mit dem die spezifischen Probleme und Rahmenbedingungen an den beteiligten Programmschulen berücksichtigt werden können (Beeth et al., 2003; Prenzel, 2000; Prenzel, 2001). Wenn man den skizzierten Ansatz an den Schulen umsetzen möchte, müssen die Ausgangslagen vor Ort (insbesondere diejenigen, die sich auf den mathematisch-naturwissenschaftlichen Bereich beziehen) berücksichtigt werden. Für eine breite Implementation und auch Dissemination der Programmkonzeption ist deshalb zu klären, wie dieser Ansatz bei unterschiedlichen Ausgangslagen auf Seiten der Schulen und der Schülerschaft umgesetzt werden kann. Gerade die Analyse von Eingangsvoraussetzungen und damit die genaue Beleuchtung von Unterschieden zwischen beteiligten Schulen werden bei der Einführung von Innovationen häufig wenig beachtet.

Eingangsvoraussetzungen sollten jedoch aus mehrerlei Gründen berücksichtigt werden. So wird erstens Implementationsprojekten vielfach entgegengehalten, dass die teilnehmenden Schulen nicht repräsentativ seien. Mögliche positive Ergebnisse von Innovationsprogrammen laufen dann Gefahr, abgewehrt zu werden, weil eine Verbreitung der Ansätze auf „normale“ Schulen nicht gewährleistet sei. Insbesondere dann, wenn es sich beispielsweise um eine Auswahl besonders leistungsstarker Programmschulen handelt, wird eine mögliche Übertragbarkeit des Ansatzes auf weniger leistungsstarke Schulen in Frage gestellt. Zweitens kann die Wirkung einer Innovation, also beispielsweise eine Veränderung der Motivation und Leistung der Schülerinnen und Schüler, allein in Bezug auf die jeweilige Ausgangslage festgestellt werden. Nur in Relation zu einer systematischen Analyse der Ausgangslage lässt sich letztlich Implementationsforschung betreiben, bei der untersucht werden soll, welche Voraussetzungen sich günstig oder ungünstig auf die Implementation eines neuen Ansatzes auswirken.

Der folgende Beitrag berichtet deshalb über Ausgangsbedingungen für die Umsetzung des Programms an den beteiligten Schulen. Im Vordergrund steht die Frage nach der Repräsentativität der beteiligten Schulen. Wichtige Merkmale auf Schulebene wurden über Befragungen der Schülerinnen und Schüler erhoben. Die Aggregation dieser Daten auf Schulebene erlaubt Aussagen über die Implementationsbedingungen an den Schulen. Als Ausgangsbedingungen betrachtet werden hier u.a. das Interesse, motivationale Orientierungen sowie die Kompetenzen der Schülerinnen und Schüler im mathematisch-naturwissenschaftlichen Bereich. Außerdem werden der familiäre Hintergrund der Schülerinnen und Schüler und das Schulklima erhoben. Im Rahmen der Implementationsforschung stellt sich dabei die Frage, wie sich die Unterschiede in diesen Merkmalen zwischen beteiligten Schulen bei der Umsetzung des Programms bemerkbar machen. Als Refe-

renz dient eine für die Bundesrepublik Deutschland repräsentative Stichprobe von Schulen, die nicht am Programm beteiligt waren.

## *1. Implementationsfragen als Teil der Begleitforschung zum BLK-Programm SINUS*

Die Begleitforschung zum BLK-Programm „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“ bezieht sich auf mehrere Aspekte: (a) Schulinterne Evaluation und Qualitätssicherung. (b) (Externe) Formative Evaluation, (c) (Externe) Summative Evaluation und (d) Implementation. Im Folgenden werden die drei erstgenannten Forschungsstränge kurz dargelegt (Prenzel, 2000), um anschließend die Implementationsforschung in den Rahmen des Evaluationskonzepts einzuordnen, die den Fokus dieses Artikels darstellt.

*Schulinterne Evaluation und Qualitätssicherung.* Ein Ziel des Programms besteht darin, die problembezogene Unterrichtsentwicklung anhand von Modulen schulintern mit Prozessen der Qualitätssicherung zu kombinieren (Prenzel, 2000). Die Schulen verständigen sich auf gemeinsam zu verfolgende Ziele und setzen einfache Evaluationsmethoden ein, um Effekte der eigenen Arbeit zu erfassen (z.B. kurze Schülerfragebogen, Formen der kollegialen Evaluation, Hospitationen, o.ä.).

*(Externe) Formative Evaluation.* Die prozessbegleitende Evaluation bezieht sich auf die Optimierung des Unterstützungsangebotes für beteiligte Lehrkräfte und darauf, wie dieses genutzt wird. Außerdem soll die Frage beantwortet werden, wie die Lehrkräfte mit dem Ansatz des SINUS-Programms zurechtkommen und inwieweit die Arbeit an den Schulen in Gang gekommen ist. Zu zwei Messzeitpunkten wurden hierzu die beteiligten Lehrkräfte befragt (Ostermeier, 2003; Prenzel, Ostermeier, Bahr & Hammann, 2000).

*(Externe) Summative Evaluation.* Bei der summativen Evaluation dienen die Ziele bzw. die Ansprüche des Programms als Kriterien für die Erfolgskontrolle. Mit dem Modellversuchsprogramm SINUS wird eine Verbesserung des mathematisch-naturwissenschaftlichen Unterrichts, des mathematisch-naturwissenschaftlichen Kompetenzniveaus der Schülerinnen und Schüler und des fach- und sachbezogenen Interesses angestrebt (Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung, 1997a). An den Schulen sollen weiterhin Verfahren der Qualitätssicherung fest eingerichtet sein und konsequent betrieben werden. Entsprechende Entwicklungen sollten sich aufgrund der kollegialen Zusammenarbeit in den Fachgruppen nicht nur bei einzelnen Lehrkräften und in deren Unterricht niederschlagen, sondern den mathematisch-naturwissenschaftlichen Unterricht an der jeweiligen Schule insgesamt betreffen.

Zur Überprüfung, inwieweit bis zum Ende des Programms (und im Sinne nachhaltiger Effekte sogar darüber hinaus) Fortschritte in Richtung der an-

gestrebten Ziele an den beteiligten Schulen festzustellen sind, wurden zwei groß angelegte Erhebungen an den SINUS-Schulen durchgeführt. Die erste Erhebung diente der Einschätzung der Eingangsvoraussetzungen zu Beginn der Arbeit auf Schulebene (Ostermeier, Carstensen, Geiser & Prenzel, 2002). Die zweite Erhebung soll eine Beurteilung der Entwicklungen im Laufe des Programms ermöglichen. Die Zielgruppe der Erhebungen waren jeweils Schülerinnen und Schüler der neunten Jahrgangsstufe sowie Schulleitungen und (bei der zweiten Erhebung) auch beteiligte, sowie nicht am Programm beteiligte Lehrkräfte.

*Implementationsforschung.* Über die Bestimmung eines Bezugspunkts (Baseline) zur summativen Untersuchung von Entwicklungen hinaus gewinnen die Daten der oben erwähnten Eingangserhebung vor allem im Kontext mit Fragestellungen der Implementationsforschung an Relevanz. In diesem Rahmen stellt sich die Frage, wie der Ansatz des SINUS-Programms an den Schulen umgesetzt wird und welche Bedingungen die Implementation behindern oder fördern (Euler & Sloane, 1998; Geijssel, Slegers, Berg & Keltchermans, 2001; Gräsel & Parchmann, 2004; Knapp, 1997; Reinmann-Rothmeier & Mandl, 1998). Entsprechende Fragen können auf der Basis der Grundinformationen aus den eben skizzierten Erhebungen bearbeitet werden. So können einerseits die Daten aus den formativen Erhebungen zu Rate gezogen werden, um zu untersuchen, welche Unterstützungen und Maßnahmen geeignet sind, die Arbeit an den Schulen anzuregen oder über Schwierigkeiten hinweg zu helfen. Daneben stellt sich die Frage, unter welchen Ausgangsbedingungen auf der Ebene der beteiligten Schulen die Implementation des SINUS-Ansatzes gelingt und unter welchen Bedingungen mit Problemen zu rechnen ist. Dieser Fragenkomplex kann mit Hilfe der Daten der oben beschriebenen Eingangserhebung bearbeitet werden und steht im Zentrum des vorliegenden Beitrags.

*Fragestellungen.* Das BLK-Programm SINUS beruht auf einem Implementationsansatz, bei dem Lehrkräfte problemorientiert und ausgehend von Modulen ihren eigenen Unterricht weiterentwickeln. Die Umsetzung des Qualitätsentwicklungsansatzes hängt dabei entscheidend von den vor Ort vorliegenden Bedingungen ab.

Aus der Sicht der Implementationsforschung gilt es zu klären, inwieweit theoretisch bedeutsame Bedingungen die Umsetzung des Programms beeinflussen und welche dieser Bedingungen sich förderlich oder hinderlich auswirken. Um Implementationsforschung betreiben zu können, müssen Informationen über relevante Bedingungsgebiete zur Verfügung stehen. Diese Informationen müssen die Ausgangslage an den beteiligten Schulen beschreiben und daher schon vor Einsetzen einer Programmwirkung vorliegen. Sie betreffen auf den mathematisch-naturwissenschaftlichen Bereich bezogene motivationale (Interesse, Selbstkonzept) und kognitive (Vorwissen) Ausgangsbedingungen auf Schulebene. Auch allgemeine Einschätzun-



gen der Schülerschaft zur Schule (z.B. Schulklima) sind bedeutsame Bedingungsfaktoren. Der hohe Zusammenhang zwischen den Ergebnissen in den Leistungstests und dem familiären Hintergrund der Schülerinnen und Schüler wurde in der PISA 2000 Studie deutlich (Baumert & Schümer, 2001). Es ist daher angezeigt, auch bei der Analyse der Ausgangsbedingungen solche Bedingungen zu überprüfen. Daneben bezieht sich eine besondere Frage auf die Verbreitung (Dissemination) der Innovation auf weitere Schulen. Geklärt werden muss, ob eine Verallgemeinerung der Implementation und damit zusammenhängend deren förderlicher und hinderlicher Bedingungen möglich erscheint. Mit dem vorliegenden Beitrag soll deshalb die Frage beantwortet werden, ob die Schulen, die am Programm beteiligt sind, eine besondere, beispielsweise überdurchschnittlich leistungsstarke, Auswahl darstellen oder, ob es sich um eine Gruppe durchschnittlicher Schulen handelt.

Der Beitrag untersucht gezielt die auf den mathematisch-naturwissenschaftlichen Bereich bezogenen Bedingungen an den beteiligten Programm-Schulen. Es stellt sich dabei die Frage, ob die SINUS-Schulen eine besondere Selektion ausgehend von den unterrichts- und schulbezogenen Ausgangslagen darstellen. Im weiteren Verlauf werden die folgenden vier Bereiche systematisch untersucht:

- Familiärer Hintergrund der Schülerinnen und Schüler
- Motivationale Orientierungen und Präferenzen im mathematisch-naturwissenschaftlichen Bereich
- Schulklima
- Kompetenzen in Mathematik und Naturwissenschaften

## *2. Erhebung der Ausgangslage an den Programm-Schulen: Design und Methode*

Für die Erhebung der Ausgangslage an den beteiligten Schulen wurde ein Untersuchungsdesign gewählt, das über die Auswahl von Erhebungsinstrumenten einen systematischen Vergleich mit repräsentativen Schultichproben gestattet. Die PISA-Studie bot die Chance, das Leistungsniveau der BLK-Schulen im Jahr 2000 vergleichend einzuordnen. Für diese PISA-Erhebung wurden von deutschen Expertengruppen zusätzliche Mathematik- und Naturwissenschaftstests entwickelt, die auf die Altersstufe der 15-Jährigen und auf die neunte Jahrgangsstufe zugeschnitten sind (Klieme, Neubrand & Lüdtke, 2001; Prenzel, Carstensen, Rost & Senkbeil, 2002; Prenzel, Rost, Senkbeil, Häußler & Klopp, 2001). Diese nationalen Instrumente wurden auch im Rahmen der Begleitforschung an den SINUS-Schulen eingesetzt. Mit der Erhebung der Ausgangslage im Jahr 2000 sollte das mathematisch-naturwissenschaftliche Kompetenzniveau und Interesse auf der Ebene der Schule zu Beginn der Arbeit im Programm festgestellt werden. Aus der Sicht des Programms ist das Unterrichts- und Kompetenz-

niveau an der jeweiligen Schule der entscheidende Ausgangs- und Bezugspunkt. Das Ziel des Modellversuchsprogramms beschränkt sich ja nicht darauf, den Unterricht einiger engagierter Lehrkräfte weiterzuentwickeln. Vielmehr sollen die professionelle Kooperation und die Qualitätsentwicklung in den Fachgruppen in Gang gesetzt werden, um pädagogisch und didaktisch wirkungsvolle Unterrichtsansätze an der Schule umzusetzen und zu verbreiten. Auch allgemein schulbezogene Variablen wie beispielsweise Schulklima oder Hintergrundvariablen (familiärer Hintergrund) wurden auf einer repräsentativen Basis erfasst.

Bedingt durch die Zeitschiene von PISA 2000 konnte die Eingangserhebung erst mit einer gewissen Verzögerung gegenüber dem Programmstart im Herbst 1998 durchgeführt werden. Hier stellt sich die Frage, wie man den Messzeitpunkt im Frühjahr 2000 interpretiert. Zum Untersuchungszeitpunkt hatten sich bereits feste Schulsets und Arbeitsgruppen engagierter Lehrkräfte an den BLK-Schulen gebildet. Typischerweise waren die beteiligten Lehrkräfte an den Schulen in der Anfangsphase damit befasst, ihre Ziele und Arbeitsschwerpunkte im Programm festzulegen. An den Schulen hatte man bis zum Erhebungszeitpunkt begonnen, Qualitätssicherung mit Hilfe ausgewählter Module und Klassenstufen zu betreiben. Es kann jedoch nicht davon ausgegangen werden, dass sich zu diesem Zeitpunkt Programmaktivitäten bereits auf der Ebene der gesamten Schule und im Unterricht ausgewirkt haben. Die Arbeiten an den Schulen haben also zu diesem Zeitpunkt noch nicht flächendeckend den Mathematik- und Naturwissenschaftsunterricht an der Einzelschule erreicht. Gleichzeitig setzt die Eingangserhebung aber mit der Testung des gesamten neunten Jahrgangs bewusst auf Schulebene an, so dass die Daten in der Tat das Ausgangsniveau der Schule widerspiegeln.

## **2.1 Erhebungsverfahren**

Die Eingangserhebung betraf verschiedene Merkmalsbereiche auf Seiten der Schülerinnen und Schüler. Hierzu zählen die mathematische und naturwissenschaftliche Kompetenz, motivationale Konstrukte (fachbezogenes Selbstkonzept und Interesse) sowie Freizeitpräferenzen und schul- bzw. fachrelevante Einstellungen. Die Erhebungsinstrumente wurden unter Rückgriff auf Verfahren der PISA-Studie zusammengestellt. Als Testverfahren wurden Mathematik- und Naturwissenschaftsaufgaben der nationalen PISA-Erweiterung verwendet. Bei den Fragebögen wurden erprobte Motivations- bzw. Selbstkonzeptskalen verwendet, die auch bei PISA eingesetzt wurden. Es wurde eine Selektion mathematik- und naturwissenschaftsrelevanter Fragen des Schülerfragebogens vorgenommen.

Der Schülerfragebogen enthielt keine expliziten Fragen zu Sozialschichtindikatoren oder soziodemographischen Variablen des Elternhauses. Es mussten Vorgaben des Datenschutzes eingehalten werden, die es uns bei der Eingangserhebung nicht gestatteten, Fragen zum sozioökonomischen Status und zur Ausbildung der Eltern zu erheben.

Tabelle 1: Skalen, auf die im vorliegenden Beitrag zurückgegriffen wird  
(vgl. Kunter et al., 2002)

<i>Bereich / Skala</i>	<i>Anz. items</i>	<i><math>\alpha</math></i>	<i>Beispielitem</i>
<i>Familiärer Hintergrund</i>			
Besitz (Ausstattung des Haushalts) <sup>1</sup>	7	.62	Gibt es bei dir zu Hause eine Geschirrspülmaschine?
Besitz an Kulturgütern (kulturelle Ressourcen) <sup>1</sup>	5	.69	Gibt es bei dir zu Hause klassische Literatur (z. B. von Goethe)?
Elterliche Unterstützung bei den Hausaufgaben <sup>1</sup>	3	.63	Wie oft unterstützen dich die folgenden Personen bei den Hausaufgaben oder anderen Arbeit für die Schule, z. B. ... deine Mutter?
<i>Motivationale Orientierungen und Präferenzen</i>			
Interesse (Mathematik) <sup>1,3</sup>	3	.79	Mathematik ist mir persönlich wichtig.
Selbstkonzept (Mathematik) <sup>1,4</sup>	3	.90	Mathematik ist eines meiner besten Fächer.
<i>Schulklima</i>			
Positive Schüler-Lehrerbeziehung (Interesse des Lehrers am Lernen und der Person) <sup>2,5</sup>	6	.81	Unsere Lehrer/innen reden mit uns, wenn uns etwas nicht gefällt.
Negative Schüler-Schüler-Beziehung (Konkurrenz und soziale Desintegration) <sup>2,6</sup>	5	.65	In unserer Klasse versucht jede/r besser zu sein als der/die andere.
Allgemeine Schulfriedenheit <sup>2,7</sup>	3	.65	Ich fühle mich in unserer Schule gut aufgehoben.

1 PISA 2000 (Internationaler Schülerfragebogen) (Kunter et al., 2002; OECD, 2001)

2 PISA 2000 (Nationaler Schülerfragebogen) (vgl. Kunter et al., 2002)

3 Der Skala liegt die Interessenskala zugrunde, die sowohl in der BIJU-Studie (Baumert, Gruehn, Heyn, Köller & Schnabel, 1997) als auch in TIMSS/II und III verwendet wurde (Schiefele, 1991; Schiefele, 1996)

4 Kurzfassung des Self Description Questionnaire (SDQ) von Marsh (1990)

5 vgl. Weishaupt, Zedler & Fickermann (1998)

6 vgl. Helmke & Dreher (1979, S. 159f.), Tillmann, Holler-Nowitzki, Holtappels, Meier und Popp (1999, S. 356 f.)

7 Die Skala wurde von Helmke & Dreher (1979, S. 159f.) auf der Grundlage von Fend (1998) entwickelt; Die verwendete Version stammt von Tillmann, Holler-Nowitzki, Holtappels, Meier & Popp (1999, S. 356 f.)

## 2.2 Bestimmung der Leistungswerte

In den Leistungstests der Eingangserhebung wurden analog zur PISA-Studie 2000 neun verschiedene Testhefte verwendet. In einem Rotationsdesign wurden Testaufgaben systematisch auf die Testhefte verteilt, wobei jede Testperson nur einen Teil der Aufgaben bearbeitete (Baumert et al.,

2002). Mit Hilfe von Verfahren der Item Response Theorie (IRT) (Rost, 2004) werden Leistungswerte für jede Schülerin und jeden Schüler unter Berücksichtigung der tatsächlich bearbeiteten Aufgaben und deren Schwierigkeiten berechnet.

Um die Leistungswerte aus den beiden Untersuchungen (SINUS vs. PISA/ PISA-E) miteinander vergleichen zu können, wurden die latenten Skalen der Eingangserhebung an den latenten Skalen der PISA-E-Studie verankert (Davier, 2002). Für die gemeinsamen Skalen von SINUS und PISA/E in den nationalen Mathematik- und Naturwissenschaftstests wurde der Mittelwert für den bundesdeutschen Leistungsdurchschnitt auf 100 Punkte festgelegt. Die Standardabweichung wurde auf 20 Punkte festgelegt.

### 2.3 Durchführung der Erhebung

Die Eingangserhebung war zunächst an allen an SINUS beteiligten Schulen vorgesehen. Im Rahmen der Stichprobenziehung der internationalen PISA-Stichproben sowie der nationalen deutschen Erweiterungsstichprobe (PISA-E) wurden jedoch 37 SINUS-Schulen gezogen, 35 davon nahmen an der PISA-Testung teil. An der BLK-SINUS-Erhebung beteiligten sich insgesamt 136 BLK-Schulen. Um für die SINUS-Schulen zuverlässige Leistungskennwerte zu erhalten, wurde ein gesamter Jahrgang (hier: 9. Jahrgang) getestet.

### 2.4 Beschreibung der Stichproben

Ein bedeutsames Merkmal, dass mit Schulleistung in Verbindung steht, ist die Schulform. Deshalb besteht ein erster Schritt darin, die Verteilung der Schüler auf Schulformen in den beiden Stichproben zu vergleichen. Tabelle 2 ist zu entnehmen, dass die Verteilung der getesteten Schülerinnen und Schüler in der Eingangserhebung deutlich von der Verteilung der getesteten PISA-Schüler abweicht. Im Vergleich zur national repräsentativen Stichprobe besuchen größere Anteile der SINUS-Schüler ein Gymnasium oder eine Integrierte Gesamtschule. Die SINUS-Stichprobe enthält im Vergleich zur national repräsentativen Stichprobe etwas weniger Realschüler. Auffallend ist, dass Schülerinnen und Schüler der Hauptschule in der BLK-Stichprobe deutlich unterrepräsentiert sind.

Tabelle 2: Verteilung der getesteten Schülerinnen und Schüler auf Schulformen

<i>Schulform</i>	<i>SINUS (Prozent)</i>	<i>PISA (Prozent)</i>
Hauptschule	5.3	22.6
Schule mit mehreren Bildungsgängen	11.7	9.7
Realschule	20.2	28.1
Integrierte Gesamtschule	18.3	10.1
Gymnasium	44.5	29.6

Diese Ungleichverteilung auf Schulformen in der BLK-Stichprobe muss freilich bei weiteren Auswertungen berücksichtigt werden. Daher werden weitere Ergebnisse in Kombination mit der Schulform berichtet.

Bei den Stichproben handelt es sich jeweils um Schülerinnen und Schüler der 9. Klasse, die im Schnitt zwischen 15 und 16 Jahre alt sind (SINUS: M: 15,75; SD: 0,82; PISA-E: M: 15,81; SD: 0,74). Die Stichproben unterscheiden sich im mittleren Alter nicht signifikant voneinander.

### 3. Ergebnisse

Die Ergebnisse der SINUS-Schulen werden stets in Relation zur national repräsentativen Stichprobe (PISA-E-Schulen) dargestellt. Für den Vergleich wurden für die Schülerskalen zweifaktorielle Varianzanalysen gerechnet. Neben dem Faktor Stichprobe mit den Ausprägungen SINUS und PISA-E wurde der Faktor Schulform mit fünf Stufen (HS: Hauptschulen, MBG: Schulen mit mehreren Bildungsgängen, RS: Realschulen, IGS: Integrierte Gesamtschulen, GYM: Gymnasien) berücksichtigt. Die Ergebnisse der Mittelwertvergleiche werden für die Bereiche familiärer Hintergrund, motivationale Variablen und Schulklima getrennt berichtet. Aufgrund der großen Anzahl an befragten Schülerinnen und Schülern werden auch relativ kleine Mittelwertsunterschiede als statistisch signifikant ausgewiesen. Ein Blick auf die Effektstärken, die auch jeweils in den Tabellen dargestellt werden, ist daher unumgänglich.

#### 3.1 Familiärer Hintergrund

Tabelle 3 zeigt die Ergebnisse der zweifaktoriellen Varianzanalyse mit mehreren gezielt ausgewählten Variablen des familiären Hintergrunds. Als Faktoren wurden einerseits die Schulform, andererseits die Stichprobe (SINUS vs. PISA-E) verwendet.

Tabelle 3: Ergebnisse der Mittelwertvergleiche mit den Faktoren Schulform (SF) und Stichprobe (ST) (SINUS vs. PISA-E) mit Skalen zum familiären Hintergrund als abhängigen Variablen

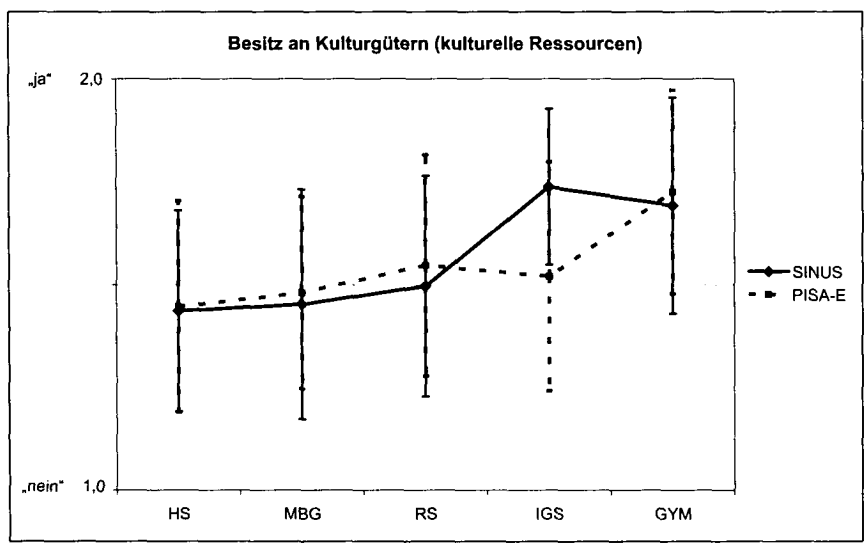
Skala	Faktor	F	Hyp. df	Fehler df <sup>d</sup>	p	$\eta^2$
Besitz	SF	20,96	4	6418	<.01	0,013
(Ausstattung des	SP	5,90	1	6418	<.05	0,001
Haushalts)	SF * SP	0,89	4	6418	ns	0,001
Besitz an Kulturgütern	SF	135,44	4	7412	<.01	0,068
(Kulturelle	SP	1,76	1	7412	ns	0,000
Ressourcen)	SF * SP	5,85	4	7412	<.01	0,003
Elterliche	SF	11,11	4	9726	<.01	0,005
Unterstützung bei	SP	36,93	1	9726	<.01	0,004
den Hausaufgaben	SF * SP	0,48	4	9726	ns	0,000

Zunächst soll überprüft werden, ob die beiden Stichproben (SINUS, PISA-E) hinsichtlich des familiären Hintergrunds vergleichbar sind. Betrachtet man die Skala „Besitz (Ausstattung des Haushalts)“, so liegen die Angaben der SINUS-Schüler im Schnitt (M: 1,58; SD: 0,20; 1 = „nein“, 2 = „ja“) unter denen der PISA-E-Schüler (M: 1,59; SD: 0,18) (Tabelle 3). Obwohl die-

ser Unterschied statistisch signifikant ausfällt, ist dieser Haupteffekt mit einer Varianzaufklärung von unter einem Prozent nicht als praktisch bedeutsam zu werten. Die Stichproben unterscheiden sich statistisch nicht in der Skala „Besitz an Kulturgütern“. Dagegen liegen die SINUS-Schüler (M: 2,01; SD: 0,99) mit ihren Angaben zur elterlichen Unterstützung bei Hausaufgaben (1 = „nie oder fast nie“ bis 5 = „mehrmals in der Woche“) unter denen ihrer gleichaltrigen Mitschüler aus der PISA-E-Stichprobe (M: 2,21; SD: 1,04). Auch hier können jedoch die Unterschiede bei einem  $\eta^2 = 0,004$  nicht als praktisch bedeutsam angesehen werden.

Überprüft wurden auch die Unterschiede zwischen den Schulformen. Dabei zeigen sich in allen vier berichteten Skalen des familiären Hintergrunds signifikante Unterschiede zwischen den Schulformen. Allerdings zeigen sich bei zwei der Skalen - Besitz (Ausstattung des Haushalts), elterliche Unterstützung bei den Hausaufgaben - vernachlässigbar kleine Effektstärken mit einer Varianzaufklärung von einem oder unter einem Prozent. Lediglich der Bereich der kulturellen Ressourcen in der Familie weist signifikante Unterschiede auf, bei einer gleichzeitig etwas höheren Varianzaufklärung durch den Faktor Schulform von sieben Prozent. Dabei fallen die Angaben der Schülerinnen und Schüler aus Gymnasien höher aus, als diejenigen der anderen Schulformen (Abbildung 1).

Abbildung 1: Mittelwerte und Standardabweichungen der Skala „Besitz an Kulturgütern (kulturelle Ressourcen)“ unterteilt nach Stichprobe und Schulform<sup>2</sup> (Beispielitem: „Gibt es bei dir zu Hause klassische Literatur (z. B. von Goethe)?“; 1 = „nein“, 2 = „ja.“)



Die Skala „Besitz an Kulturgütern“ weist gleichzeitig den einzigen signifikanten Interaktionseffekt zwischen Schulform und Stichprobe auf. Dieser wird durch die Schulform Gesamtschule bedingt. Wie Abbildung 1 zeigt,

berichten die Gesamtschüler der SINUS-Schulen über bessere kulturelle Ressourcen in ihren Familien. Allerdings ist auch hier die Stärke dieses Effekts ( $\eta^2 = 0,003$ ) nur sehr gering.

### 3.2 Motivationale Orientierungen und Präferenzen

Die Motivation und das Interesse der Schülerinnen und Schüler im mathematisch-naturwissenschaftlichen Bereich ist eine relevante Ausgangsbedingung für Unterricht. Das BLK-Programm zielt auf eine Förderung affektiv-motivationaler Variablen. Für die Steuerung und Beurteilung der Implementation sind deshalb solide Informationen über das Ausgangsniveau des Interesses und weiterer motivationaler Variablen im mathematisch-naturwissenschaftlichen Bereich bedeutsam. Die Befunde der Eingangserhebung für das Interesse und das Selbstkonzept der Schülerinnen und Schüler werden in Tabelle 4 gezeigt. Dargestellt wird in den Vergleichen lediglich der Bereich Mathematik. Die entsprechenden Skalen für den Bereich Naturwissenschaften wurden bei PISA nur einer sehr kleinen Stichprobe vorgelegt, die keinen repräsentativen Vergleich gestattet.

Tabelle 4: Ergebnisse der Mittelwertsvergleiche mit den Faktoren Schulform (SF) und Stichprobe (SP) (SINUS vs. PISA-E) mit Skalen zum Interesse und Selbstkonzept (Mathematik) als abhängigen Variablen.

Skala	Faktor	F	Hyp. df	Fehler df'	p	$\eta^2$
Besitz	SF	20,96	4	6418	<.01	0,013
(Ausstattung des	SP	5,90	1	6418	<.05	0,001
Haushalts)	SF * SP	0,89	4	6418	ns	0,001
Besitz an Kulturgütern	SF	135,44	4	7412	<.01	0,068
(Kulturelle	SP	1,76	1	7412	ns	0,000
Ressourcen)	SF * SP	5,85	4	7412	<.01	0,003
Elterliche	SF	11,11	4	9726	<.01	0,005
Unterstützung	SP	36,93	1	9726	<.01	0,004
bei den Hausaufgaben	SF * SP	0,48	4	9726	ns	0,000

Sowohl bei der abhängigen Variable „Interesse an Mathematik“, als auch bei der abhängigen Variable „Selbstkonzept Mathematik“ zeigen sich sehr signifikante Haupteffekte im Vergleich zwischen den Stichproben. Dabei liegen die Angaben der Schülerinnen und Schüler aus SINUS-Schulen sowohl im Bereich des Mathematik-Interesses (M: 2,24; SD: 0,82), als auch im Bereich Selbstkonzept (M: 2,42; SD: 0,96) unter denen der PISA-E-Schülerinnen und Schüler (Interesse Mathematik, M: 2,38; SD: 0,82; Selbstkonzept Mathematik, M: 2,49, SD: 0,95). Wiederum müssen die geringen Effektstärken ( $\eta^2 = 0,002 / 0,001$ ) bedacht werden. Diese gestatten es nicht, die signifikanten Unterschiede als praktisch bedeutsam zu interpretieren. Gleiches gilt für den Haupteffekt Schulform, der für die abhängige Variable „Interesse Mathematik“ zu beobachten ist. Der Anteil aufgeklärter Varianz beträgt hier lediglich 0,3 Prozent. Beide Interaktionen fallen nicht signifikant aus.

### 3.3 Schulklima

Nicht ausgeschlossen werden kann, dass die Umsetzung des Programms neben den eher auf den Fachunterricht bezogenen Bedingungen bedeutsam vom vorhandenen Klima an der Schule abhängt. Für den Bereich Schulklima wurden drei Skalen ausgewählt, die die Schüler-Lehrerbeziehung, die Beziehungen der Schülerinnen und Schüler untereinander sowie die allgemeine Schulzufriedenheit betreffen. Tabelle 5 zeigt die Ergebnisse der zweifaktoriellen Varianzanalysen mit den entsprechenden Skalen als abhängigen Variablen.

Tabelle 5: Ergebnisse der Mittelwertsvergleiche mit den Faktoren Schulform (SF) und Stichprobe (SK) (SINUS vs. PISA-E) mit Skalen zum Schulklima als abhängigen Variablen

Skala	Faktor	F	Hyp. df	Fehler df <sup>1</sup>	p	$\eta^2$
Positive	SF	13,66	4	3495	<.01	0,015
Schüler-Lehrer-	SP	0,08	1	3495	ns	0,000
Beziehung	SF * SP	2,76	3	3495	<.05	0,002
Negative	SF	1,72	4	8226	ns	0,001
Schüler- Schüler-	SP	0,35	1	8226	ns	0,000
Beziehung	SF * SP	2,08	4	8226	ns	0,001
Allgemeine	SF	20,13	4	7429	<.01	0,011
Schul-	SP	241,75	1	7429	<.01	0,032
zufriedenheit	SF * SP	11,90	4	7429	<.01	0,006

Für die Skalen Schüler-Lehrerbeziehungen und Schüler-Schüler-Beziehungen zeigen sich keine statistisch abgesicherten Unterschiede zwischen den Stichproben (SINUS, PISA-E). Dagegen liegen die Angaben der SINUS-Schülerinnen und -Schüler im Bereich der allgemeinen Schulzufriedenheit (M: 3,33; SD: 0,97) signifikant über denen der PISA-E Stichprobe (M: 2,77; SD: 0,61; Antwortformat: 1=„stimmt gar nicht“ bis 5=„stimmt ganz genau“). Auch wenn sich für die Schulzufriedenheit als abhängige Variable ein signifikanter Haupteffekt für den Faktor Schulform sowie ein signifikanter Interaktionseffekt zwischen Schulform und Stichprobe nachweisen lässt, fallen dennoch die Effektstärken gering aus. Beim Haupteffekt des Faktors Stichprobe ist die Effektstärke zwar auch sehr gering. Die Varianzaufklärung von drei Prozent liegt jedoch verglichen mit den übrigen beobachteten Werten etwas höher.

Betrachtet man die Skala positive Schüler-Lehrerbeziehung als abhängige Variable zeigt sich ein sehr signifikanter Haupteffekt des Faktors Schulform und eine signifikante Interaktion. Beide Effekte sind jedoch gering und praktisch nicht bedeutsam.

### 3.4 Kompetenzen im mathematisch-naturwissenschaftlichen Bereich

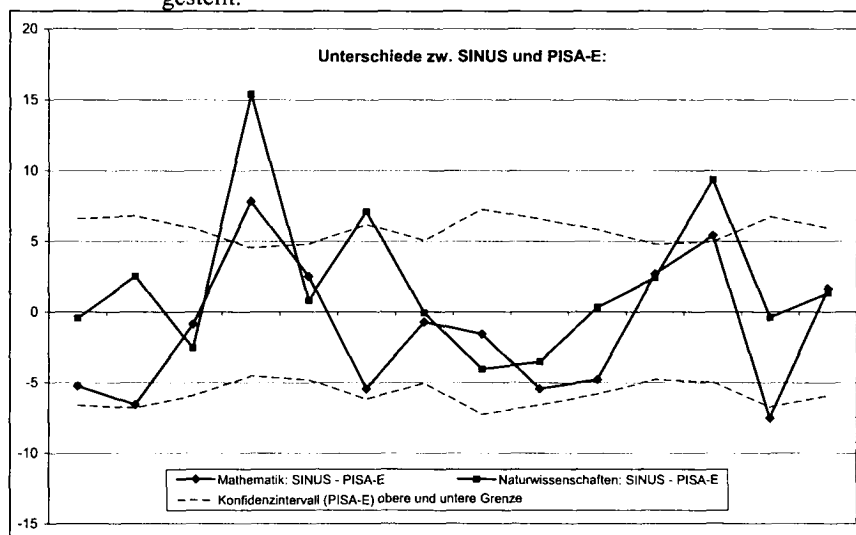
Für den nun folgenden Vergleich der Leistungen, in PISA-Schulen und in SINUS-Schulen zu Beginn des Programms, wurden Differenzwerte berich-



tet. Dabei wurde jeweils der durchschnittliche Punktwert der PISA-Stichprobe vom Punktwert der SINUS-Schulen abgezogen, so dass Werte über Null einen Leistungsvorsprung von SINUS-Schulen anzeigen. Der Vergleich erfolgt in Gruppen, nach Ländern (der Bundesrepublik Deutschland) und Schulformen differenziert. Die folgende Abbildung zeigt eine Profildarstellung dieser Differenzwerte getrennt für Mathematik und Naturwissenschaften.

Abbildung 2 gibt Differenzen der Leistungskennwerte exemplarisch für Gymnasien wieder. Unterteilt wird nach den Ländern, in denen Gymnasien am Programm teilnahmen. In dieser Abbildung werden ferner Grenzen von Konfidenzintervallen für die Differenzen angezeigt, die allerdings für Mathematik und Naturwissenschaften gemittelt dargestellt sind. Sie dienen nur dazu, die Größenordnung der Vertrauensintervalle anzugeben und sind nicht im Sinne einer Signifikanzprüfung zu interpretieren. Anders verhält es sich bei den darauf folgenden Abbildungen (Abbildung 3 und Abbildung 4). Die dort angegebenen Konfidenzintervalle beruhen auf einer getrennten Berechnung der Standardschätzfehler für die SINUS- und PISA-E-Schulen. Deshalb können dort die Konfidenzintervalle im Sinne eines Signifikanztests interpretiert werden.

Abbildung 2: Leistungsvergleich zwischen SINUS-Schulen und PISA-E-Schulen. Angezeigt werden Punktdifferenzen (SINUS minus PISA-E) in Profildarstellung differenziert nach Ländern (Schulform: Gymnasium). Zusätzlich werden Konfidenzintervalle für die Punktdifferenzen dargestellt.

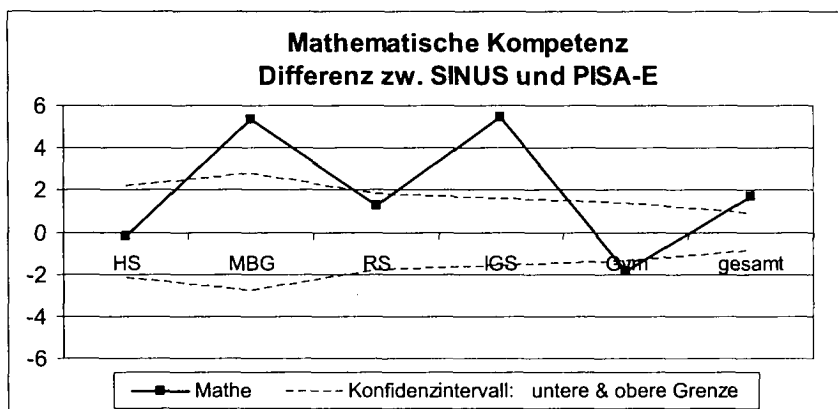


Die Profildarstellung verdeutlicht, dass SINUS-Schulen bezogen auf die Leistungswerte nicht generell über bzw. unter der national repräsentativen PISA-Stichprobe liegen. Allerdings zeigen sich nennenswerte Unterschiede

zwischen einzelnen Schulen. Die differenzielle Betrachtung der Gruppen, die durch die Kombination aus Schulform und Land gebildet werden, weist insgesamt auf eine starke Heterogenität der SINUS-Schulen hin. Neben Schulen, die sich nicht bedeutsam von der national repräsentativen PISA-E Stichprobe unterscheiden, finden sich andere, die deutlich von der mittleren Leistung der PISA-E-Schulen abweichen. So gibt es auch Schulen, deren Leistungen deutlich über den entsprechenden PISA-E-Schulen derselben Schulform im eigenen Land liegen. Aber auch umgekehrt finden sich SINUS-Schulen, die im Vergleich zu Schulen derselben Schulform im selben Land niedrigere Leistungswerte aufweisen.

In den folgenden Abbildungen (Abbildung 3 für Mathematik, Abbildung 4 für Naturwissenschaften) wird der Leistungsvergleich zwischen allen SINUS-Schulen und der PISA-E-Stichprobe dargestellt. Die Profile zeigen wiederum die Punktdifferenzen zwischen SINUS-Schulen und PISA-E-Schulen an, wobei Werte über Null einen Punktvorsprung der SINUS-Schulen bedeuten. Die Graphiken zeigen Differenzen nach Schulformen. Außerdem werden in der Darstellung Konfidenzintervalle abgebildet. Die Standardschätzfehler wurden dabei getrennt aus den Daten von SINUS und PISA-E errechnet. Insgesamt ist es daher möglich, die Grenzen der Konfidenzintervalle im Sinne eines Signifikanztests zu interpretieren.

Abbildung 3: Leistungsvergleich zwischen SINUS-Schulen und PISA-E-Schulen. Angezeigt werden Punktdifferenzen (SINUS minus PISA-E) für die mathematische Kompetenz in Profildarstellung differenziert nach Schulformen<sup>2</sup>

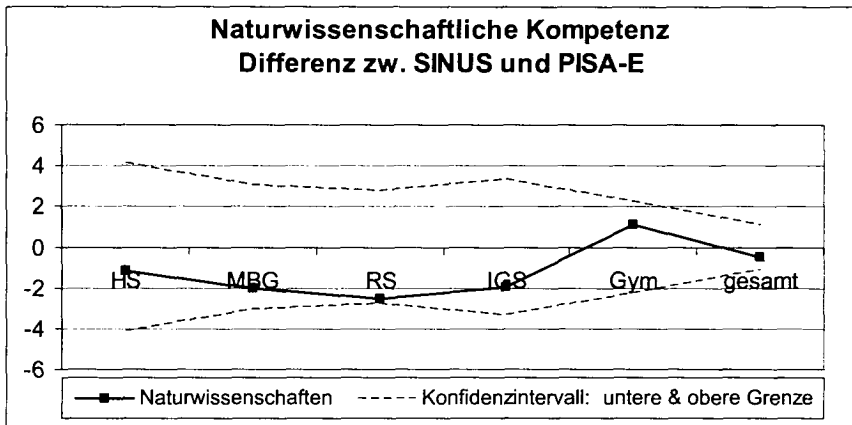


Ausgehend davon ergeben sich je nach betrachtetem Kompetenzbereich unterschiedliche Bilder. Für die mathematische Kompetenz liegen die SINUS-Schulen der Schulen mit mehreren Bildungsgängen, sowie die Integrierten Gesamtschulen über den Leistungen der PISA-E-Schulen gleicher Schulform (Abbildung 3). SINUS-Gymnasien liegen im Durchschnitt etwas unter den PISA-E-Gymnasien. Die Gesamtleistung der SINUS-Schulen liegt im

Bereich der Mathematik insgesamt signifikant über der Leistung der PISA-E-Schulen.

Bei der naturwissenschaftlichen Kompetenz zeigt sich ein anderes Bild (Abbildung 4). Es liegen die Differenzen zwischen SINUS- und PISA-E-Schulen innerhalb der Grenzen des Konfidenzintervalls, die Abweichungen bewegen sich also im Bereich des Zufalls.

Abbildung 4: Leistungsvergleich zwischen SINUS-Schulen und PISA-E-Schulen. Angezeigt werden Punktdifferenzen (SINUS minus PISA-E) für die naturwissenschaftliche Kompetenz in Profildarstellung differenziert nach Schulformen<sup>2</sup>



#### 4. Diskussion

Insgesamt erbringt die SINUS-Eingangserhebung wichtige Antworten auf die Frage, ob sich das Programm auf eine Selektion besonderer Schulen bezieht. Es zeigen sich keine systematischen Leistungsvorsprünge oder -rückstände wenn man sämtliche an der Eingangserhebung beteiligten Programm-Schulen betrachtet. Bemerkenswert ist jedoch die Leistungsheterogenität der Schulen. Einige Schulen ragen offensichtlich in ihrer Leistung über die national repräsentative Stichprobe hinaus. Viele BLK-Schulen sind schließlich hinsichtlich der Leistungsdifferenzen unauffällig

##### 4.1 Bedeutung der Ergebnisse für die Implementationsforschung

Die Implementationsforschung zum BLK-Programm SINUS soll klären, welche theoretisch bedeutsamen Merkmale oder Merkmalsmuster beteiligter Schulen sich förderlich oder hinderlich auf die Umsetzung des Ansatzes an den Schulen auswirken. Dieser Frage kann nur nachgegangen werden, wenn systematische Daten zur Ausgangslage zu Beginn der Arbeit vorliegen. Diese Daten stehen mit der Eingangserhebung zur Verfügung.

Neben den Hinweisen auf die generelle Vergleichbarkeit der SINUS-Schulen, die unter der nächsten Überschrift diskutiert werden, enthalten die

Daten wichtige Informationen über spezifische Implementationsvoraussetzungen an den SINUS-Schulen. In dieser Hinsicht sind weniger die allgemeinen Unterschiede und Vergleichbarkeiten auf der Ebene der gesamten Stichprobe, als die Variabilität zwischen den einzelnen Schulen von Bedeutung. Im vorliegenden Beitrag wurden zwar keine Unterschiede in den Daten einzelner SINUS-Schulen berichtet. Dennoch weisen die Darstellungen nach Schulgruppen (z.B. Schulformen und Länder) auf die Variabilität innerhalb der Gruppe der Programm-Schulen hin. Bei den Skalen des familiären Hintergrunds beispielsweise finden sich nennenswerte Effektstärken im Bereich kultureller Ressourcen und zwar bezogen auf die Unterschiede zwischen den Schulformen.

Sichtbar wird die Variabilität innerhalb der Programm-Schulen vor allem dann, wenn man die Leistungsdifferenzen gruppiert nach Schulformen innerhalb der Länder betrachtet. SINUS-Schulen sind demnach mit unterschiedlichen Voraussetzungen auf Seiten ihrer Schülerschaft in die Programmarbeit eingetreten. In weiteren Forschungsarbeiten können beispielsweise Kompetenzentwicklungen betrachtet werden in Abhängigkeit der spezifisch von den Schulen gewählten Module. Ferner kann die Umsetzung des Programms auch im Hinblick auf die beteiligten Schulformen untersucht werden.

## **4.2 Vergleichbarkeit der Implementationsbedingungen:**

### **Bedeutung der Ergebnisse für die Verbreitung des Programms**

Es wurde auch der Frage nachgegangen, ob sich die Programmschulen auf eine besondere Auswahl von Schulen beziehen, oder ob die beteiligten Schulen in theoretisch bedeutsamen Variablen als repräsentativ angesehen werden können. Dieser Frage wurde mit fachunterrichtsbezogenen (Interesse, Selbstkonzept, Kompetenzen im mathematisch-naturwissenschaftlichen Bereich), hintergrundsbezogenen (familiärer Hintergrund) und allgemein schulbezogenen (Schulklima) Variablen gezielt nachgegangen. Als Referenz diente eine national repräsentative Stichprobe der PISA-E-Schulen.

Betrachtet man die einzelnen in diesem Beitrag berichteten Bereiche, so zeigen sich unterschiedliche Bilder. Bei den Skalen zum familiären Hintergrund zeigen sich zwischen den beiden Stichproben zwar in zwei Fällen signifikante Unterschiede. So fallen die Angaben der SINUS-Stichprobe bei den Skalen Besitz (Ausstattung des Haushalts) und elterliche Hausaufgabenunterstützung geringer aus als in der PISA-Stichprobe. Die Abweichungen sind jedoch aufgrund der sehr geringen Effektstärken praktisch nicht bedeutsam. Der einzige, ernst zu nehmende Unterschied ergibt sich beim Vergleich zwischen den Schulformen. Hier zeigt sich erwartungsgemäß, dass die Schülerinnen und Schüler aus Gymnasien, zum Teil auch aus integrierten Gesamtschulen, höhere Angaben zum kulturellen Kapital machen, als beispielsweise die Schülerinnen und Schüler aus den Hauptschulen. Dieses Ergebnis weist in Richtung des PISA-Befundes, dass in Deutschland die Sozialschichtzugehörigkeit stärker als gedacht mit Bil-

derungsergebnissen der Schülerinnen und Schüler zusammenhängt (Baumert & Schümer, 2001). Auch für die Implementation eines unterrichtsbezogenen Qualitätsentwicklungsansatzes, der letztlich auf die Steigerung von Kompetenzen auf Seiten der Schülerschaft abzielt, sind deshalb die untersuchten Hintergrundmerkmale bedeutsam.

Im Rahmen der Eingangserhebung konnte auch das Interesse und das Selbstkonzept der Schülerinnen und Schüler im Bereich Mathematik zwischen den beiden Stichproben verglichen werden. Trotz signifikanter Mittelwertsunterschiede können die Differenzen wiederum aufgrund der äußerst geringen Effektstärken nicht als praktisch bedeutsam gewertet werden. Die SINUS-Schulen sind in dieser Hinsicht mit der repräsentativen Stichprobe vergleichbar.

Für den Bereich Schulklima wurden drei Skalen berichtet. Während sich die Stichproben bei den Skalen Schüler-Lehrerbeziehung und Schüler-Schüler-Beziehung nicht signifikant unterscheiden, zeigt sich ein Effekt dieses Faktors im Bereich allgemeine Schulfriedenheit. Die Schülerinnen und Schüler der SINUS-Programmschulen scheinen also tendenziell zufriedener mit ihren Schulen zu sein als die Schülerschaft der repräsentativen PISA-Stichprobe. Allerdings liegt für diesen signifikanten Haupteffekt die Varianzaufklärung auch lediglich bei drei Prozent. Wenn man einmal von dieser niedrigen Effektstärke absieht, so dürfte die positivere Einschätzung des Schulklimas durch die Schülerinnen und Schüler grundsätzlich eine eher förderliche Bedingung zur Umsetzung des Programms an den beteiligten Schulen darstellen.

Die Ergebnisse in den Leistungstests müssen getrennt für die beiden Domänen Mathematik und Naturwissenschaften betrachtet werden. Die Vergleichbarkeit mit der repräsentativen Stichprobe scheint im Fall der Naturwissenschaften unproblematisch zu sein, da die Leistungsdifferenzen innerhalb der Grenzen der Konfidenzintervalle liegen. Unabhängig von der Schulform scheinen die SINUS-Schulen sich in ihrer durchschnittlichen Leistung also nicht von der repräsentativen Stichprobe zu unterscheiden. In dieser Hinsicht handelt es sich bei den am Programm beteiligten Schulen um eine Auswahl normaler, also weder überdurchschnittlich herausragender, noch überdurchschnittlich entwicklungsbedürftiger, Schulen. Stellt sich heraus, dass die Implementationsstrategie bei dieser Auswahl von Schulen greift, so steigen die Chancen, den Ansatz auch auf eine größere Gruppe von Schulen zu verbreiten.

Im Fall der Mathematik zeigt sich ein etwas anderes Bild. Für zwei Schulformen - Schulen mit mehreren Bildungsgängen und integrierte Gesamtschulen - liegen die durchschnittlichen Mathematikleistungen in den SINUS-Schulen signifikant über den Werten der Vergleichsstichprobe. Dies lässt vermuten, dass sich in diesen Fällen Schulen für das Programm angemeldet haben bzw. gewählt wurden, die von vornherein eine gewisse Affi-

nität zum mathematischen Bereich aufwiesen. Im Vergleich dazu liegt die Gruppe der Gymnasien in der SINUS-Stichprobe in den Mathematikleistungen signifikant unter den Werten der Vergleichsgruppe. Hier könnte es sein, dass sich Schulen für das Programm interessiert haben, die sich gezielt im mathematischen Bereich weiter entwickeln wollten.

Die Kenntnis dieser Ausgangslage hilft, unterstützende und hinderliche Bedingungen für die Umsetzung der SINUS-Konzeption zu identifizieren. So kann beispielsweise im Kontrast der beiden Domänen (Mathematik und Naturwissenschaften) untersucht werden, ob überdurchschnittliche Leistungen im Kompetenzbereich eine notwendige Bedingung für die Implementation des Ansatzes darstellen oder ob eine Umsetzung auch unter durchschnittlichen Voraussetzungen gelingen kann.

Bei der Analyse der Folgeerhebung des Jahres 2003 wird es darum gehen, Entwicklungen der SINUS-Schulen im mathematisch-naturwissenschaftlichen Kompetenzbereich aufzuzeigen. Die Folgeerhebung wurde im Jahr 2003 unter Rückgriff auf nationale PISA-Instrumente durchgeführt. Aufgrund der Anbindung an die PISA-2003-Zeitschiene werden die Auswertungen frühestens 2005 zur Verfügung stehen. Das Ziel bestand darin, Entwicklungen an den SINUS-Schulen im Hinblick auf das mathematisch-naturwissenschaftliche Interesse sowie des Kompetenzniveaus der Schülerinnen und Schüler zu untersuchen. Dabei gewinnen die hier vorgestellten Ergebnisse an Relevanz, wenn es darum geht Bezugspunkte für Entwicklungen zu bestimmen. Ferner kann die Untersuchung, ob und an welchen Stellen die Implementation des SINUS-Ansatzes gelungen ist oder nicht, aufgrund der Daten der Ausgangserhebung systematisch mit den Implementationsvoraussetzungen in Beziehung gesetzt werden.

### *Anmerkungen*

- 1 Aufgrund von Stichprobeneigenschaften variiert die so genannte effektive Stichprobengröße und damit auch die Zahl der Freiheitsgrade von Variable zu Variable. Die effektive Stichprobengröße hängt vom Grad der Abhängigkeit der beobachteten Werte innerhalb der Schulen ab. Je stärker die beob. Werte innerhalb der Schulen streuen, gemessen an der Streuung zwischen den Schulen, desto niedriger ist diese Abhängigkeit. Näheres und Details dazu werden zum Beispiel im Kap. 1.6.9 in Baumert, Artelt, Carstensen, Sibberns und Stanat (2002) berichtet.
- 2 HS: Hauptschulen, MBG: Schulen mit mehreren Bildungsgängen, RS: Realschulen, IGS: Integrierte Gesamtschulen, Gym: Gymnasien.

### *Literatur*

Baumert, J., Artelt, C., Carstensen, C. H., Sibberns, H. & Stanat, P. (2002). Untersuchungsgegenstand, Fragestellungen und technische Grundlagen der Studie. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000 - Die*

- Länder der Bundesrepublik im Vergleich* (S. 11-38). Opladen: Leske + Budrich.
- Baumert, J., Gruehn, S., Heyn, S., Köller, O. & Schnabel, K.-U. (1997). *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU). Dokumentation, Band 1. Skalen Längsschnitt I, Welle 1-4*. Berlin.
- Baumert, J. & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbe- teiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 323-407). Opladen: Leske + Budrich.
- Beeth, M., Duit, R., Prenzel, M., Ostermeier, C., Tytler, R. & Wickman, P. O. (2003). Quality development projects in science education. In D. Psillos, P. Kariotoglou, V. Tsifas, G. Fassouloupoulos, E. Hatzikraniotis & M. Kallery (Hrsg.), *Science education research in the knowledge based society* (S. 447- 457). Dordrecht: Kluwer.
- Blumenfeld, P. C., Fishman, B. J., Krajcik, J. S. & Marx, R. W. (2000). Crea- ting usable innovations in systemic reform: Scaling up technology- embedded project-based science in urban schools. *Educational Psychologist*, 35, 149-164.
- Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung. (1997a). *Gutachten zur Vorbereitung des Programms „Steigerung der Effi- zienz des mathematisch-naturwissenschaftlichen Unterrichts“*. Bonn: Bun- desministerium für Bildung, Wissenschaft, Forschung und Technologie.
- Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung. (1997b). *Neuordnung der Modellversuchsförderung*. Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie. Internet: [http://www. blk-bonn.de/modellversuche/neuordnung.htm](http://www.blk-bonn.de/modellversuche/neuordnung.htm) [21.09.2003].
- Davies, v. M. (2002). *BLK-PISA Skalierungs- und Linkstudie*. Princeton, NJ: Educational Testing Service (ETS).
- Euler, D. & Sloane, P. F. E. (1998). Implementation als Problem in der Mo- dellversuchsforschung. *Unterrichtswissenschaft*, 26, 312-326.
- Fend, H. (1998). *Qualität im Bildungswesen: Schulforschung zu Systembedin- gungen, Schulprofilen und Lehrerleistung*. Weinheim: Juventa.
- Geijssels, F., Sleegers, P., Berg, R., van den & Keltchermans, G. (2001). Condi- tions fostering the implementation of large-scale innovation programs in schools: Teachers' perspectives. *Educational Administration Quarterly*, 37, 130-166.
- Gräsel, C. & Parchmann, I. (2004). Die Entwicklung und Implementation von Konzepten situierten, selbstgesteuerten Lernens. *Zeitschrift für Erziehungs- wissenschaften*, 6 (2. Beiheft), 171-184.
- Helmke, A. & Dreher, E. (1979). *Gesamtschule und dreigliedriges Schulsystem in Nordrhein-Westfalen: Erzieherische Wirkungen und soziale Umwelt*. Pa- derborn: Schöningh.
- Klieme, E., Neubrand, M. & Lüdtke, O. (2001). Mathematische Grundbildung: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 139-190). Opladen: Leske + Budrich.

- Knapp, M. S. (1997). Between systemic reforms and the mathematics and science classroom: The dynamics of innovation, implementation, and professional learning. *Review of Educational Research*, 67, 227-266.
- Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J. & Weiß, M. (2002). *PISA 2000: Dokumentation der Erhebungsinstrumente*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Marsh, H. W. (1990). *Self Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescent self-concept: An interim test manual and a research monograph*. San Antonio, TX: The Psychological Corporation.
- OECD (Hrsg.). (2001). *Knowledge and skills for life: First results from PISA 2000*. Paris: OECD.
- Ostermeier, C. (2003). *Kooperative Qualitätsentwicklung in Schulnetzwerken am Beispiel des BLK-Modellversuchsprogramms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“*. (Unveröffentlichte Dissertation) Kiel: IPN.
- Ostermeier, C., Carstensen, C. H., Geiser, H. & Prenzel, M. (2002). *Eingangserhebung im BLK-Modellversuchsprogramm „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“: Rückmeldung der Ergebnisse an die beteiligten Schulen*. (Arbeitsbericht) Kiel: IPN.
- Prenzel, M. (2000). Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts: Ein Modellversuchsprogramm von Bund und Ländern. *Unterrichtswissenschaft*, 28, 103-126.
- Prenzel, M. (2001). Das BLK-Modellversuchsprogramm „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“. In E. Klieme & J. Baumert (Hrsg.), *TIMSS: Impulse für Schule und Unterricht* (S. 59-65). Bonn: BMBF.
- Prenzel, M. & Achtenhagen, F. (2000). Innovation durch Modellversuchsprogramme. *Unterrichtswissenschaft*, 28, 98-102.
- Prenzel, M., Carstensen, C. H., Rost, J. & Senkbeil, M. (2002). Naturwissenschaftliche Grundbildung im Ländervergleich. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000 - Die Länder der Bundesrepublik Deutschland im Vergleich* (S. 129-158). Opladen: Leske + Budrich.
- Prenzel, M. & Ostermeier, C. (2003). Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts - Ein unterrichtsbezogenes Qualitätsentwicklungsprogramm. *Beiträge zur Lehrerbildung. Fachzeitschrift zu Theorie und Praxis der Aus- und Weiterbildung von Lehrerinnen und Lehrern*, 21, 265-276.
- Prenzel, M., Ostermeier, C., Bahr, S. & Hammann, M. (2000). *Befragung zur Akzeptanz im BLK-Modellversuchsprogramm „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“: Ergebnisse und Implikationen für die weitere Steuerung des Programms*. (Arbeitsbericht) Kiel: IPN.
- Prenzel, M., Rost, J., Senkbeil, M., Häußler, P. & Klopp, A. (2001). Naturwissenschaftliche Grundbildung: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat,



- K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 191-248). Opladen: Leske + Budrich.
- Reinmann-Rothmeier, G. & Mandl, H. (1998). Wenn kreative Ansätze versanden: Implementation als verkannte Aufgabe. *Unterrichtswissenschaft*, 26, 292-311.
- Rost, J. (2004). *Lehrbuch Testtheorie - Testkonstruktion*. Bern: Huber.
- Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, 26, 299-323.
- Schiefele, U. (1996). *Motivation und Lernen mit Texten*. Göttingen: Hogreve.
- Schön, D. A. (1987). *Educating the reflective practitioner. Toward a new design for teaching and learning in the professions*. San Francisco: Jossey-Bass.
- Tillmann, K.-J., Holler-Nowitzki, B., Holtappels, H. G., Meier, U. & Popp, U. (1999). *Schülergewalt als Schulproblem*. Weinheim: Juventa.
- Weishaupt, H., Zedler, P. & Fickermann, D. (1998). *Kleine Grundschulen in Europa: Berichte aus elf europäischen Ländern*. Weinheim: Dt.-Studien-Verlag.

Anschrift der Autoren:

Christian Ostermeier, Claus H. Carstensen, Manfred Prenzel, Helmut Geiser  
 Leibniz-Institut für die Pädagogik der Naturwissenschaften  
 Olshausenstraße 62  
 24098 Kiel

Durch Beurteilung eingereichter Manuskripte und Veränderungsvorschläge haben in diesem Jahr folgende Gutachter die Herausgeber der „Unterrichtswissenschaft“ unterstützt:

K.H. Arnold, Hildesheim	M. Kerres, Duisburg-Essen
K.O. Bauer, Bochum	K. Konrad, Weingarten
R. Biehler, Kassel	A. Lohaus, Merburg
S. Blömeke, Berlin	H. Neber, München
M. Cierpka, Heidelberg	M. Neubrand, Oldenburg
M. Clausen, Mannheim	H. Niegemann, Erfurt
D. Euler, St. Gallen	G. Pätzold, Dortmund
M. Fölling-Albers, Regensburg	C. Pauli, Zürich
H.F. Friedrich, Tübingen	H.G. Rossbach, Bamberg
M. Henninger, Weingarten	D. Rost, Marburg
F.W. Hesse, Tübingen	M. v. Saldern, Lüneburg
A. Huber, Weingarten	R. Stark, Saarbrücken
L.J. Issing, Berlin	S. Willberg, Lüneburg

An dieser Stelle sei allen Gutachtern für ihre Unterstützung gedankt.

---

## Berichtigung

---

Sehr geehrte Leserin, sehr geehrter Leser,

im letzten Heft (3/2004) wurde auf S. 227 leider eine falsche Tabelle 4 abgedruckt. Die richtige Tabelle ist im Folgenden abgedruckt:

Tabelle 4: Ergebnisse der Mittelwertsvergleiche mit den Faktoren Schulform (SF) und Stichprobe (SP: SINUS vs. PISA-E) mit Skalen zum Interesse und Selbstkonzept (Mathematik) als abhängigen Variablen.

Skala	Faktor	F	Hyp. df	Fehler df <sup>1</sup>	p	$\eta^2$
Interesse (Mathematik)	SF	9,72	4	13907	<.01	0,003
	SP	34,05	1	13907	<.01	0,002
	SF * SP	0,11	4	13907	ns	0,000
Selbstkonzept (Mathem.)	SF	1,14	4	10281	ns	0,000
	SP	8,43	1	10281	<.01	0,001
	SF * SP	1,42	4	10281	ns	0,001

Wir bitten um Ihr Verständnis.

Die Schriftleitung